Human-Computer Interaction – INTERACT '84 / B. Shackel (ed.)
Elsevier Science Publishers B.V. (North-Holland)

A COMPARISON OF SELECTION TECHNIQUES: TOUCH PANEL, MOUSE AND KEYBOARD

John Karat, James E. McDonald and Matt Anderson

International Business Machines Corporation
Austin, Texas 78758

A study was conducted testing user performance and attitudes for three types of selection devices. The subjects were tested on target selection practice tasks, and in typical computer applications using menu selection and keyboard typing. The study showed an advantage for on-screen touch panel over keyboard selection, and for keyboard selection over mouse entry. Differences between this result and those reporting an advantage of mouse selection are discussed.

1. INTRODUCTION

The process of doing work using a computer system can be broken down into three steps. The user must understand the goal of the work to be accomplished, formulate a task solution (a series of steps to accomplish a work goal), and carry out the plan. Whether the user is solving old problems with a new system, or is experienced with the computer as a problem solving tool, the final stages require relating the steps of the solution to the system in a "language" that the system will understand.

Given that humans have only limited information processing resources, one goal for system designers is to make the process of relating the user's goal to the system as easy and natural as possible. One technique has been the development of interfaces using menus from which an alternative is selected, rather than typed from memory. While menu selection on a computer has generally involved typing on a keyboard, several recent workstations have designed interfaces utilizing "more natural" selection mechanisms (e.g., pointing and voice recognition).

If we examine the process of making selections from everyday lists, we would find several common selection mechanisms. These include verbalizing the selection ("I would like a cheeseburger and fries."), pointing with the hand to a selection on a list ("Give me this one.") or possibly marking a selection with a pencil. It is interesting to note that selection in a restaurant reverts to pointing under many circumstances (e.g., unpronounceable entrees). One argument for utilizing devices such as touch panels or joysticks in a computer system is that they provide a mechanism which allows the user to point in a relatively natural fashion (i.e. with a mechanism that tracks hand movement).

The purpose of the current study was to examine three ways of selecting items in a menu driven computer system. The methods considered were selection by pointing and touching an on-screen item with the hand, selection by moving an on-screen pointer with an off-screen device, and selection by typing an identifier associated with a given display item. These methods can be thought of as representing different degrees of "naturalness" of selection.

There are likely to be tradeoffs in the consideration of selection mechanisms. Though the mechanics of typing have been studied extensively (e.g., [3]), relatively little is known about the limitations of other selection mechanisms in general computer dialogue use. It seems intuitive that some pointing devices are more natural than keyboards for some applications (e.g., graphics), but little is known about the merits of these devices in applications which currently require the use of a keyboard for input (e.g., text tasks). Selection by pointing has been evaluated in a number of recent studies which carry out movement of an on-screen pointer through manipulation of off-screen devices, such as keyboard cursor keys, joysticks, tracker balls and mice [1]. These studies generally find the mouse to be the best off-screen pointer manipulation device, and the mouse was used in this study.

The current experiment was designed to address some points not considered in reported work. First, the tasks commonly used in pointer studies were target selection tasks (tasks in which a subject points to a target randomly positioned by the system on a VDT). The work reported here uses a menu system where users were asked to solve problems using the system. We would like to verify the conclusions of target experiments in a more realistic computer dialogue. Next, most previous studies have not included touch panels. While there is some recent work which has included evaluations of touch screens [7], this work is confined to target selection tasks. Additionally, previous studies have failed to evaluate the most common form of menu selection, the typing of an identifier associated with an item. Finally, we included a situation in which menu selection was only part of the dialogue. Some of our experimental tasks involved typing information on the system keyboard.

1.1 A Simplified Theory of Selection

While the structure, the terminology used, and a host of other factors undoubtedly affect the use of a menu system, the current work focuses on the selection process. This is done to provide a manageable problem domain for applied cognitive psychology. We treat the selection process as an event consisting of several components. This technique has been found useful in modeling human cognitive behavior [4, 2].

There are many well learned "natural" selection procedures. If someone makes a menu selection in normal human interaction, it is not generally necessary to instruct them in how to do it (although someone did at one time). For human-computer interactions, a menu can help the user to decide <u>what</u> to do, and <u>how</u> to do it. The user may have to be instructed in a "new" selection method (i.e. telling them that the machine can't hear them or doesn't know where their finger is pointing). For a new user there may be cognitive costs associated with learning the new method. If the selection method is not a simple extension of well learned procedures, the user must devote attention to figuring out how to convey the selection to the system. On the other hand, studies have shown that well learned methods tend to be performed very rapidly and with little cognitive requirements (e.g., [5]).

With a "natural" menu selection the user:

> Carries out the decision process,
> then indicates the choice.

For a "new" selection process, the user must:

> Carry out the decision process,
> figure out how to indicate the choice,
> then indicate the choice.

With experience, the additional steps required to figure out how to indicate the choice become well learned themselves. That is, the procedure for indicating a choice becomes automatic, and the individual does not have to devote additional resources to it. The exact mechanisms for procedure automation are not considered here (see [6], for a more detailed discussion), only the distinction between new and relatively well known procedures is needed. Even though people do learn new procedures, a desirable system design goal is to limit device-specific knowledge (knowledge required to operate specific system components). This makes learning through the integration of system procedures into existing knowledge easier by making computer dialogue consistent with "natural" dialogues.

This might suggest an advantage for direct pointing, but the issue is still complex. Card, Moran and Newell [2] have shown that formal task analysis can be used to aid in making design decisions. This task analysis breaks behavior down into a number of cognitive cycles and motor components, and can be used to predict operator performance. If we assume users are familiar with all three methods, it is possible to carry out such an analysis and reduce selection for each to a single thought cycle and a single physical movement. Though the distances moved by the hand may vary for the three methods, the difference would be slight, and the analysis would suggest little difference between the methods. However, it does not seem that this analysis captures all elements needed to insure "ease of use" for a range of users.

While a keystroke-level analysis might work for well learned tasks, it is not clear that it can be extended to more complex tasks. One problem with using such a model in a cognitively complex domain is estimating with any certainty the number of cognitive cycles required for an activity. The thought cycles are viewed by Card et al. [2] as corresponding to a processing cycle of about 100 msec. duration. While it might be possible to agree on the number of cognitive steps of this time order required for fairly simple well learned tasks, cognitive science has not developed sufficiently to provide detailed models for much of normal cognitive behavior. It is not clear that pointing with a mouse requires the same amount of cognitive processing as finger pointing, or if switching input devices increases processing. We could provide estimates of cognitive load for switching between input devices, but we would certainly want to verify empirically such estimates before accepting them as valid. Since a large percentage of task time might be spent in simply making selections, it is important to carefully consider the tradeoffs involved.

2. METHOD

2.1 Subjects

Twenty-six subjects (25 females and one male) were paid to participate in the study. Subjects were obtained from a temporary employment agency and ranged in age from 18 to 57 years with self reported typing rates ranging from 30 to over 100 words per minute. Forty-two percent of the subjects had obtained college degrees, fifty percent had some college but no degree, and the remaining eight percent had completed high school. Most of the subjects (83%) reported some word processing experience, but only four had any programming experience. Two subjects failed to complete the experimental tasks and were excluded from the analyses. All reported analyses are based on data obtained from the 24 subjects who successfully completed the experiment.

2.2 Apparatus and Task Environments

An IBM Personal Computer (PC) served as the experimental controller. Three independent input devices were attached to the PC (a touch panel, a mouse, and a keyboard) such that each could be used to select options from menus presented on the "task display" (a 13 inch IBM Color

Monitor). The input devices were an Elographics analog-membrane touch panel, a Mouse Systems optical mouse, and a standard IBM PC keyboard. The "instruction display" (a 12 inch IBM Monochrome Monitor) was used to present instructions, application descriptions, and individual menu tasks to the subjects. Subjects sat facing the two displays with the keyboard located between the subject and the displays throughout the experiment. The touch panel was attached to the front of the task display during all input device conditions. The mouse was located on a metal pad next to the subject's preferred-hand side of the keyboard during the mouse-device conditions, and was moved away for keyboard and touch panel conditions. The keyboard was used as a menu selection device for the keyboard-device condition, and for entering text into "fields" during all three conditions.

Two computer-based systems were developed for the experiment and simulated on the PC. The two menu task environments (applications) simulated a computer telephone aid and a personal appointment calendar. Both of the applications appeared to the subject as hierarchically organized menus from which the subjects selected options. Additionally, half of the tasks that the subjects were asked to perform required the entry of typed information into the system (such as names, comments or telephone numbers). Thus subjects performed tasks (e.g., calling someone from a directory or changing an appointment) which required selecting options with a device and typing into selected fields from the keyboard. The system responded to subject input by appearing to perform the specified actions (actual telephone connections were not made nor did an appointment data base exist).

The two applications differed in structure and complexity. The depth (number of selections needed to reach a given action) and breadth (number of options available from a given state) of the two systems' menu hierarchies were different. The number of menu selections for the required tasks ranged from one to eight for the telephone system tasks, and from three to eleven for the calendar tasks. The calendar tasks included: reviewing daily schedules, checking past and future appointments, scheduling new appointments, and modifying existing appointments. The telephone tasks involved calling persons listed in various directories and modifying those directories by changing the contents of entries, deleting entries, or creating new entries.

2.3 Experimental Design

All subjects used all three input devices to perform tasks within both applications. This resulted in a 3 (device) by 2 (application) within-subjects design. The order in which the subjects used the input devices was based on a Latin-Square, with three device orderings used (keyboard/touch panel/mouse, touch panel/mouse/keyboard, and mouse/keyboard/touch panel). Application order (telephone/calendar or calendar/telephone) was counterbalanced across subjects. This resulted in a total of six groups to which the subjects were assigned in a pseudo-random fashion. Within each condition, the order of presentation of individual tasks and applications remained constant across devices. A set of practice trials preceeded each set of tasks. During an experimental session, subjects used the first input device to perform 25 practice trials, 10 tasks in the first application, 25 additional practice trials, and then performed 10 tasks in the second application. This procedure was repeated for the two remaining input devices. In total, subjects completed two sets of practice trials (one set before each application) and two sets of application tasks (10 tasks in each task environment) for each of three devices.

2.4 Procedure

Subjects received practice with the appropriate input device prior to performing tasks within an application. Practice consisted of selecting a target "button" which appeared in random locations on the task display. Subjects were instructed to select the target as rapidly as possible while keeping selection errors to a minimum. Targets were outlined boxes (14x13 mm) with a single letter inside. On each practice trial the screen would go blank and then the target box would appear. The position of the target box and the box label ("a" through "z") were randomly selected on each trial.

Selection with the touch panel involved touching anywhere inside the target with the subject's finger. Selection with the keyboard involved typing the letter which appeared inside the target box. Selection with the mouse involved positioning the screen pointer inside the target and pressing any button on the mouse. Movements of the mouse resulted in equal movements (1:1) of the pointer on the task display. Target selection time was measured from the appearance of the target on the display to the selection of the target with the input device.

Following the 25 practice trials subjects read general task instructions along with a brief description of the upcoming application (no attempt was made to develop a complete instruction manual for the applications). The first menu panel of the appropriate application appeared on the task display and the task descriptions were presented sequentially on the instruction screen. The successful completion of one task (detected by the simulation system) initiated the presentation of the next task description and returned the application to the first menu panel. Subjects learned the telephone and calendar systems by attempting to complete the assigned tasks. No additional assistance was provided.

Subjects performed the same 20 tasks using each of the three input devices. Visually, the menus were identical for each input device condition.

A single prompt line at the bottom of the screen displayed "Select your choice." (since the tasks followed practice with the current selection method, we did not expect or observe confusion with how selections were to be made). All tasks required menu selections, and half of the tasks also required text entry into "fields". In order to enter text into fields, subjects first selected the target field on the task display. After selection the field was "activated" by the system, and visually identified by a prompt line. Subjects then typed the required information into the selected field, using backspace correction if necessary, and deactivated the field when finished by pressing the RETURN key. Task completion time was measured from the presentation of the problem on the instruction screen to completion of the entries to satisfy the goal stated in the problem.

3. RESULTS

3.1 Preference Data

Subjects were asked to indicate preferences between the input devices for the experimental tasks. After all tasks were completed, subjects were asked to answer a series of questions comparing all pairs of input devices for different task types (practice, selection only, and selection with typing), and applications. There was no significant difference between the preference ratings for the two applications (calendar and telephone), or between the practice and selection only tasks. Table 1 summarizes the results for the 24 subjects.

	Selection Only			Selection with Typing		
	Touch	Key	Mouse	Touch	Key	Mouse
First	12	11	1	7	17	0
Second	11	10	3	14	6	4
Third	1	3	20	3	1	20

Table 1. Preference order for input devices.

For all tasks the subjects preferred the keyboard or touch panel to the mouse. The keyboard and touch panel were given similar preference ratings for selection only tasks. For tasks with both selection and typing, subjects preferred the keyboard over the touch panel. The difference between stated preference based on task type indicates some desire by subjects not to switch input devices.

3.2 Performance Data - Practice Tasks

In total, each subject completed two sets of 25 practice trials for each of the three input devices. Equipment problems caused practice data from five of the subjects to be lost. For the remaining 19 subjects a 3 (device) by 2 (set) within subject analysis of variance for the mean of each set of practice trials was performed. Main effects for device ($F(2,36) = 54.9$) and set ($F(1,18) = 36.9$), and the device by set interaction ($F(2,36) = 10.2$), were all significant. Subjects were faster making selections with the touch panel and keyboard (0.8 and 1.1 secs. overall) than with the mouse (2.7 secs.). All devices showed improvement for second set over first, showing a strong practice effect. The difference between the mouse and the other devices was greater on the first set (1.0 and 1.2 for the touch screen and keyboard respectively and 3.3 secs. for the mouse) than the second set (0.7 and 1.0 secs. for touch panel and keyboard and 2.0 secs. for mouse).

3.3 Performance Data - Menu Tasks

Subjects completed the application tasks three times, once with each input device. Table 2 summarizes the total time to complete the 20 menu tasks for the three repetitions. Each block includes 10 tasks within each application. Note that subjects using one device for the first block used a different device on the second and third blocks.

	Block		
	First	Second	Third
Touch panel	1805.3	874.9	676.2
Keyboard	1862.1	991.8	808.5
Mouse	1884.5	1189.7	897.5

Table 2. Time in secs. to complete menu tasks.

There is a clear tendency to complete the task sets faster with experience. Subjects took over 30 minutes on the average to complete the first block of 20 tasks (10 calendar and 10 telephone tasks), and less than half that time to complete the tasks on the third repetition. This can be attributed to experience with the applications. Remember that the actual applications were new to the subjects, and the first block times include considerable learning times.

The times to complete the menu tasks were subjected to an analysis of variance. Presentation order of the devices was a between subjects variable with eight subjects in each of three device order groups (keyboard/touch panel/mouse, mouse/keyboard/touch panel, and touch panel/-mouse/keyboard). Each subject used three devices, two task types (selection only and selection with typing) and two applications.

Several of the main effects are not of particular interest here. While there is a significant difference between performance on the two applications ($F(1,21) = 59.9$, with the telephone task faster), the two task types ($F(1,21) = 152.7$, with selection only faster), and an interaction between application and task type performance ($F(1,21) = 48.8$), no attempt was made to control or equate these factors. The focus was on providing varied task types and applications, without studying the relative difficulty.

There is a significant difference in performance for the three input devices ($F(2,42) = 6.1$). Additional analysis shows a significant difference between the touch panel and mouse performance. No significant difference was found for comparisons of the keyboard with mouse or touch panel. Additionally there is no significant interaction between task type and device, application and device, or task by application by device (i.e. no difference in the pattern of the main effect of device for the different applications or tasks). Thus there is no difference in the device effects for different tasks or applications.

4. DISCUSSION

Both the subjective evaluations and the performance data suggest that the touch panel and the keyboard are better selection devices than the mouse. Performance measures in the applications showed an improvement of approximately 10% for touch panel over keyboard and 10% for keyboard over mouse. The difference was similar for selection only and mixed selection and typing tasks. For practice tasks, which were simple target selection tasks, the performance difference is much greater. However, it should be kept in mind that actual selection represents only a part of the application dialogue (it is estimated that selection time composed approximately 12 - 25% of the task time for the given experiment).

Several interesting considerations result from this work. Given that the subjects in this experiment were all skilled typists, and we assume also skilled pointers, it is perhaps not surprising to find a performance advantage for typed and hand pointing mechanisms. The mouse was a new device for all subjects (note that performance gains over the practice trials were greater for the mouse than for the other devices). The performance results are slightly different than those one might expect from the preference data. Even though the subjects stated a preference for keyboard entry in mixed entry tasks, the performance data shows an advantage for touch panel input.

Given the research reports of advantages of the mouse as an input device, the failure of the mouse in the current work is interesting. Three possible sources for this might be mentioned. The first possible factor is typing skill. Since the this study used only skilled typists, it may be that this biased the study in favor of the well known device (the keyboard). Additional studies should test both skilled and non-skilled typists in a controlled fashion. The second factor is sex of the subjects. This study involved mostly female subjects, and there has been considerable investigation of sex differences in spatial processing tasks. The manipulation of the mouse to control the positioning of a screen pointer can be considered as a spatial task. Further work should be done to examine male and female subjects. The third factor is amount of practice. Even unskilled typists are far more familiar with the keyboard than they are with the mouse. While limitations on our experiment prohibit us from developing "mouse experts", we feel that additional information could be obtained by extending the amount of practice with each selection device.

It would be difficult to account for these findings within the framework suggested by Card et al. [2]. One way to account for the results of this study is to suggest that keyboard menu entry and mouse selection required additional cognitive processing compared to finger pointing. This would be consistent with the notion that touch selection is a highly automated skill for most humans, and that other techniques are less well learned. We do not suggest that this study renders previous findings invalid, only that more attention needs to be given to the nature of the dialogues for which the device is to be used, and the skills of the users. At the very least we must conclude that "unnatural" devices such as a keyboard, can in some circumstances both be preferred and lead to better performance than "natural" pointing devices such as a mouse. This study is part of a series being conducted to examine the issues raised.

REFERENCES:

1. Card, S. K., English, W. K., and Burr, B. J. (1978). Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT. Ergonomics, 21, 601-613.

2. Card, S. K., Moran, T. P., and Newell, A. (1983). The Psychology of Human-Computer Interaction. Hillsdale, N.J.: Erlbaum.

3. Cooper, W. E. (1983). Cognitive Aspects of Skilled Typewriting. New York, N.Y.: Springer-Verlag.

4. Newell, A. & Simon, H. A. (1972). Human Problem Solving, Hillsdale, N.J.: Prentice Hall.

5. Shiffrin, R. M., and Dumais, S. T. (1981). The development of automatism. In J. R. Anderson, (Ed.), Cognitive Skills and Their Acquisition, 111-140, Hillsdale, N.J.: Erlbaum.

6. Shiffrin, R. M., and Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. Psychological Review, 84, 127-190.

7. Whitfield, D., Ball, R. G., and Bird, J. M. (1983). Some comparisons of on-display and off-display touch input devices for interaction with computer generated displays. Ergonomics, 26, 1033-1054.